Excercise Tutorial 1

17th January 2025

1 Data Structures

Use the raw sales data as given below of a logistics company that tracks shipments, their costs, and the regions they serve. The task is to preprocess this data for analysis, which involves:

1. Compute average value quantity and cost of each item.

2. Compute total cost of shipments by region and total quantity shipped per region.

3. Categorizing the shipment based on cost.

   - 'Low' : less than $ 1500
   - 'Medium' : between $ 1500 to $ 3000
   - 'High' : more than $ 3000

4. Count the frequency of each item shipped across all regions.

5. Identify the most frequent item in each region.

data = [{"shipmentId": "S001", "item": "Laptop", "quantity": 5, "cost": 1400, "region": "North America"},
{"shipmentId": "S002", "item": "Phone", "quantity": 10, "cost": None, "region": "Europe"},
{"shipmentId": "S003", "item": "Tablet", "quantity": 3, "cost": 2000, "region": "Asia"},
{"shipmentId": "S004", "item": "Phone", "quantity": 2, "cost": 3000, "region": "Europe"},
{"shipmentId": "S005", "item": "Monitor", "quantity": 1, "cost": 8000, "region": "North America"},
{"shipmentId": "S006", "item": "Laptop", "quantity": 4, "cost": 4000, "region": "Asia"},
{"shipmentId": "S007", "item": "Monitor", "quantity": 2, "cost": 1500, "region": "North America"},
{"shipmentId": "S008", "item": "Phone", "quantity": 5, "cost": 2500, "region": "Asia"},

{"shipmentId": "S009", "item": "Tablet", "quantity": 8, "cost": 2000, "region": "Europe"},
{"shipmentId": "S010", "item": "Monitor", "quantity": 4, "cost": None, "region": "Asia"}
]

2 Numpy

Construct a grayscale image I using a 2D NumPy aaray of size $H \times W$ where each element in I is a pixel intensity value ranging from 0 (black) to 255 (white). Perform the following image processing and matrix-related operations.

1. Compute the **mean**, **median**, and **standard deviation** of pixel values in the image.

2. Create a binary image B where $B[i, j] = 1$ if $I[i, j] \geq T$, otherwise $B[i, j] = 0$. T is a user-defined threshold.

3. Apply a 3×3 mean filter to the image to remove noise. Replace each pixel value with the average of its 3×3 neighborhood (ignoring boundary pixels).

4. Identify bright spots in the image: A pixel is a bright spot if its value is > 200, and it is the maximum in its 3×3 neighborhood. Return their coordinates.